

# LEAST SQUARES SHADOWING METHOD FOR SENSITIVITY ANALYSIS OF DIFFERENTIAL EQUATIONS

MARIO CHATER\*, ANGXIU NI\*, PATRICK J. BLONIGAN\*, AND QIQI WANG\*

**Abstract.** For a parameterized hyperbolic system  $\frac{du}{dt} = f(u, s)$  the derivative of the ergodic average  $\langle J \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T J(u(t), s)$  to the parameter  $s$  can be computed via the Least Squares Shadowing algorithm (LSS). We assume that the system is ergodic which means that  $\langle J \rangle$  depends only on  $s$  (not on the initial condition of the hyperbolic system). After discretizing this continuous system using a fixed timestep, the algorithm solves a constrained least squares problem and, from the solution to this problem, computes the desired derivative  $\frac{d\langle J \rangle}{ds}$ . The purpose of this paper is to prove that the value given by the LSS algorithm approaches the exact derivative when the discretization timestep goes to 0 and the timespan used to formulate the least squares problem grows to infinity.

**Key words.** Sensitivity analysis, Dynamical systems, Chaos, Uniform hyperbolicity, Ergodicity, Least squares shadowing

**AMS subject classifications.** 34A34, 34D30, 37A99, 37D20, 37D45, 37N99, 46N99, 65P99

**1. Introduction.** Consider the differential equation parameterized by  $s \in \mathbb{R}$  and governing  $u(t) \in \mathbb{R}^m$  :

$$\begin{cases} \frac{du}{dt} = f(u, s) \\ u(0) = u_0 \end{cases} \quad u_0 \in \mathbb{R}^m \quad (1.1)$$

The differential equation is assumed to be uniformly hyperbolic (details in section 3). We are also given a  $C^1$  cost function  $J(u, s) : \mathbb{R}^m \times \mathbb{R} \rightarrow \mathbb{R}$  and assume that the system is *ergodic*, i.e., the infinite time average :

$$\langle J \rangle(s) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T J(u(t), s) dt \quad (1.2)$$

depends on  $s$  but does not depend on the initial condition  $u(0)$ . The differentiability of  $\langle J \rangle$  with respect to  $s$  has been proven by Ruelle [1]. Obtaining an estimation of  $\frac{d\langle J \rangle}{ds}$  is crucial in many computational and engineering problems. Indeed, many applications involve simulations of nonlinear dynamical systems that exhibit a chaotic behavior. For instance, chaos can be encountered in the following fields : climate and weather prediction [2], turbulent combustion simulation [3], nuclear reactor physics [4], plasma dynamics in fusion [5] and multi-body problems [6]. The quantities of interest are often time averages or expected values of some cost function  $J$ . Estimating the derivative of  $\langle J \rangle$  is particularly valuable in:

- **Numerical optimization.** The derivative of  $\langle J \rangle$  with respect to a design parameter  $s$  is used by gradient-based algorithms in order to efficiently optimize the design parameters in high dimensional design spaces (see [7]).
- **Uncertainty quantification.** The derivative of  $\langle J \rangle$  with respect to a parameter  $s$  gives a useful information for assessing the error and uncertainty in the computed  $\langle J \rangle$  (see [8]).

---

\*Department of Aeronautics and Astronautics, MIT, 77 Mass Ave, Cambridge, MA 02139, USA, email addresses: chaterm@mit.edu (Mario Chater), niangxiu@mit.edu (Angxiu Ni), blonigan@mit.edu (Patrick J. Blonigan), qiqi@mit.edu (Qiqi Wang)

For example, we could obtain a useful information about the impact of mankind on the climate by computing the derivative of the long time averaged global mean temperature to the amount of anthropogenic emissions ([9] shows how sensitivity analysis is used in climate studies). In the simulation of a turbulent airflow over an aircraft, estimating the derivative of the long time averaged drag to a shape design parameter is of extreme importance for engineers allowing them to improve their design [10]. It has been shown that in many of these practical examples, the quantities of interest exhibit ergodic properties, popularly known as *chaotic hypothesis* [11], [12].

When it comes to computing  $\frac{d\langle J \rangle}{ds}$ , conventional methods based on linearizing the initial value problem (1.1) become ill-conditioned when the system is chaotic. They compute derivatives that are orders of magnitude too large and the error grows exponentially larger as the simulation runs longer [13],[14]. This failure is due to the so-called *butterfly effect* and the explanation has been published by Lea et al. [13]. Some algorithms have been developed to overcome this failure. Lea et al. proposed the ensemble adjoint method which applies the adjoint method to many random trajectories, then averages the computed derivatives [13], [15]. However, the algorithm is computationally expensive even for small dynamical system such as Lorenz's one. Based on the fluctuation dissipation theorem, Abramov and Majda provided an algorithm that successfully computes the desired derivative [16]. Nonetheless, this algorithm assumes the dynamical system to have an equilibrium distribution similar to the Gaussian distribution, an assumption often violated in very dissipative systems. Recent work by Cooper and Haynes has alleviated this limitation by introducing a nonparametric method to estimate the equilibrium distribution [17]. More methods have been developed to compute  $\frac{d\langle J \rangle}{ds}$ , in particular the *Least Squares Shadowing (LSS)* algorithm which computes it by solving a constrained least squares problem [14]. The big advantage of this method is its simplicity since the least squares problem can easily be formulated and efficiently solved as a linear system. Compared to the previously presented methods, LSS is less sensitive to the dimension of the dynamical system and doesn't require any explicit knowledge about its steady-state distribution in phase space.

This paper provides a theoretical foundation for LSS by proving that it gives a useful estimation of  $\frac{d\langle J \rangle}{ds}$  when the dynamical system is a uniformly hyperbolic flow. Compared to the discrete case (uniformly hyperbolic map) for which we already have a proof of convergence [18], the continuous case is more difficult to deal with due to the apparition of the *neutral subspace* (details in section 3). However, it is very important to treat the continuous case since most applications and real-life problems require a continuous description of the physics and involve differential equations.

In the next section, the mathematical formulation of convergence is introduced as well as theorem LSS which will be proved in the following sections. Section 3 presents the concept of uniform hyperbolicity for the readers who are not familiar with the subject. Section 4 points out the new behaviour and properties that come with continuous dynamical systems (in opposition to discrete maps). Section 5 defines the shadowing direction and proves its existence as well as uniqueness. Section 6 shows that the derivative of  $\langle J \rangle$  can be computed using the shadowing direction and bounds the upper error. Section 7 then demonstrates that the least squares problem gives a good approximation of the shadowing direction. Finally, section 8 uses all the previous

results and concludes the proof of theorem LSS by showing that the estimation error vanishes as the least squares problem increases in size.

**2. Discretizing the problem and LSS theorem.** In order to obtain an algorithm of practical relevance, a discrete version of the above problem should be formulated. First, we replace the differential equation (1.1) parametrized by  $s$  by a family of operators: let  $\varphi_s(u, h) : \mathbb{R}^m \times \mathbb{R}^{+*} \rightarrow \mathbb{R}^m$  be the family of maps parametrized by  $s \in \mathbb{R}$  such that it is a "discretization" of the differential equation using a uniform time stepsize of  $h$ . In other words, if  $\{u(t), t \in (-\infty, +\infty)\}$  is a trajectory satisfying the initial differential equation (1.1) for a particular  $s \in \mathbb{R}$ , we have :

$$\varphi_s(u(t), h) = u(t + h) \quad \forall t \in \mathbb{R}$$

In this case,  $\varphi_s(\cdot, h)$  corresponds to a perfect numerical integration scheme with a stepsize of  $h$ . We ask for the differential equation to be "smooth" enough so that the maps  $\varphi(\cdot, \cdot)$  are  $C^2$ .

Then, assuming that all trajectories  $\{u(t), t \in \mathbb{R}\}$  belong to a compact set  $\Lambda$  and that  $s$  also lies in a compact set  $S \subset \mathbb{R}$ , we can approximate  $\frac{1}{T} \int_0^T J(u(t), s) dt$  with a Riemann sum and have the following bound on the integration error:

$$\left| \frac{1}{T} \left( \int_0^T J(u(t), s) dt - \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} h J(\varphi_s(u(0), ih), s) \right) \right| \leq h \sup_{s \in S} (\|DJ(\cdot, s)\|_\Lambda^\infty) \sup_{s \in S} (\|f(\cdot, s)\|_\Lambda^\infty) \quad (2.1)$$

where  $\varphi_s(u(0), 0) = u(0)$  and  $\|f(\cdot, s)\|_\Lambda^\infty$ ,  $\|DJ(\cdot, s)\|_\Lambda^\infty$  are the infinite norms of  $f(\cdot, s)$  and the derivative of  $J(\cdot, s)$  with respect to the first variable on the compact set  $\Lambda$  respectively. Since the bound doesn't depend on  $T$ , we finally have :

$$\left| \langle J \rangle(s) - \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} h J(\varphi_s(u(0), ih), s) \right| = O(h) \quad (2.2)$$

To simplify the expressions, we introduce the new notation  $\varphi_s(u(0), ih) = u_i^{s\{h, T\}}$  or even  $\varphi_s(u(0), ih) = u_i^{\{h, T\}} = u_i^s = u_i$  depending on which parameter is fixed and when there is no ambiguity.

For a sequence  $\{u_i, i = 1, \dots, \lfloor \frac{T}{h} \rfloor\}$  satisfying  $u_{i+1} = \varphi_s(u_i, h)$ , the *Least Squares Shadowing* method attempts to compute the derivative  $\frac{d\langle J \rangle}{ds}$  via

**THEOREM 2.1 (THEOREM LSS).** *Under ergodicity and hyperbolicity assumptions,*

$$\begin{aligned} \frac{d\langle J \rangle}{ds}(s) &= \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, T\}} + \partial_s J(u_i, s) + (n_i^{\{h, T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \\ &= \lim_{T \rightarrow \infty} \lim_{h \rightarrow 0} \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, T\}} + \partial_s J(u_i, s) + (n_i^{\{h, T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \blacksquare \end{aligned}$$

where  $(v_i^{\{h,T\}}, \eta_i^{\{h,T\}}) \in \mathbb{R}^m \times \mathbb{R}$ ,  $i = 1, \dots, [\frac{T}{h}]$  is the solution to the constrained least squares problem :

$$\begin{aligned} \min \sum_{i=1}^{[\frac{T}{h}]} (\|v_i^{\{h,T\}}\|^2 + \alpha(\eta_i^{\{h,T\}})^2) \\ \text{s.t. } v_{i+1}^{\{h,T\}} = (D\varphi_s(u_i, h))v_i^{\{h,T\}} + \partial_s\varphi_s(u_i, h) + h\eta_i^{\{h,T\}}\partial_h\varphi_s(u_i, h), \end{aligned} \quad (2.3)$$

where  $\alpha$  is any positive constant and  $\|\cdot\|$  is the Euclidean norm in  $\mathbb{R}^m$ .

Here the linearized operators are defined as :

$$\begin{aligned} (DJ(u, s))v &:= (D_v J)(u, s) := \lim_{\epsilon \rightarrow 0} \frac{J(u + \epsilon v, s) - J(u, s)}{\epsilon} \\ (D\varphi_s(u, h))v &:= (D_v \varphi_s)(u, h) := \lim_{\epsilon \rightarrow 0} \frac{\varphi_s(u + \epsilon v, h) - \varphi_s(u, h)}{\epsilon} \\ \partial_s J(u, s) &:= \lim_{\epsilon \rightarrow 0} \frac{J(u, s + \epsilon) - J(u, s)}{\epsilon} \\ \partial_s \varphi_s(u, h) &:= \lim_{\epsilon \rightarrow 0} \frac{\varphi_{s+\epsilon}(u, h) - \varphi_s(u, h)}{\epsilon} \\ \partial_h \varphi_s(u, h) &:= \lim_{\epsilon \rightarrow 0} \frac{\varphi_s(u, h + \epsilon) - \varphi_s(u, h)}{\epsilon} = f(\varphi_s(u, h)) \end{aligned} \quad (2.4)$$

$(DJ), (\partial_s J), (D\varphi_s), (\partial_s \varphi_s)$  and  $(\partial_h \varphi_s)$  are a  $1 \times m$  vector, a scalar, an  $m \times m$  matrix, an  $m \times 1$  vector and an  $m \times 1$  vector, respectively, representing the partial derivatives.

**3. Uniform hyperbolicity.** Let us now proceed to the presentation of the uniform hyperbolicity property : we say that the dynamical system (1.1) has a compact, global, uniformly hyperbolic attractor  $\Lambda \subset \mathbb{R}^m$  at  $s$  if for all  $t$  the map  $\varphi_s(\cdot, t)$  satisfies:

1. For all  $u_0 \in \mathbb{R}^m$ ,  $\text{dist}(\Lambda, u(t)) \xrightarrow{t \rightarrow \infty} 0$  where  $\text{dist}$  is the Euclidean distance in  $\mathbb{R}^m$ .
2. There is a  $C \in (0, +\infty)$  and  $\lambda \in (0, 1)$ , such that for all  $u \in \Lambda$ , there is a splitting of  $\mathbb{R}^m$  representing the space of perturbations around  $u$  :

$$\mathbb{R}^m = V^+(u) \oplus V^-(u) \oplus V^0(u) \quad (3.1)$$

where the subspaces are :

- $V^+(u) := \{v \in \mathbb{R}^m / \|(D\varphi_s(u, t))v\| \leq C\lambda^{-t}\|v\|, \forall t < 0\}$  is the *unstable subspace* at  $u$ ,
- $V^-(u) := \{v \in \mathbb{R}^m / \|(D\varphi_s(u, t))v\| \leq C\lambda^t\|v\|, \forall t > 0\}$  is the *stable subspace* at  $u$ .
- $V^0(u) := \{\alpha f(u, s), \forall \alpha \in \mathbb{R}\} = \{\alpha \partial_h \varphi_s(u_{i-1}, h), \forall \alpha \in \mathbb{R}\}$  is the *neutral subspace* at  $u$ .

$V^-(u), V^+(u)$  and  $V^0(u)$  are all continuous with respect to  $u$ .

If  $r = r^+ + r^- + r^0$  with  $r^+ \in V^+(u)$ ,  $r^- \in V^-(u)$ ,  $r^0 \in V^0(u)$  and  $u \in \Lambda$ , the continuity of the three subspaces and the compactity of  $\Lambda$  implies that:

$$\inf_{u, r^+, r^-, r^0} \frac{\|r^+ + r^- + r^0\|}{\max(\|r^+\|, \|r^-\|, \|r^0\|)} = \beta > 0 \quad (3.2)$$

This is because if  $\beta = 0$ , then by the continuity of  $V^+(u)$ ,  $V^-(u)$ ,  $V^0(u)$  and the compactness of  $\{(u, r^+, r^-, r^0) \in \Lambda \times \mathbb{R}^m \times \mathbb{R}^m / \max(\|r^+\|, \|r^-\|, \|r^0\|) = 1\}$ , there must be a  $(u, r^+, r^-, r^0)$  such that  $\max(\|r^+\|, \|r^-\|, \|r^0\|) = 1$  and  $r^+ + r^- + r^0 = 0$  which contradicts assumption (3.1). Thus:

$$\|r^+\| \leq \frac{\|r\|}{\beta} \quad \|r^-\| \leq \frac{\|r\|}{\beta} \quad \|r^0\| \leq \frac{\|r\|}{\beta} \quad (3.3)$$

The *stable*, *unstable* and *neutral subspaces* are also *invariant* under the differential of the map  $\varphi_s(\cdot, h)$ , i.e., if  $u_{i+1} = \varphi_s(u_i, h)$  and  $v_{i+1} = (D\varphi_s(u_i, h))v_i$ , then

$$\begin{cases} v_i \in V^+(u_i) & \Leftrightarrow & v_{i+1} \in V^+(u_{i+1}), \\ v_i \in V^-(u_i) & \Leftrightarrow & v_{i+1} \in V^-(u_{i+1}), \\ v_i \in V^0(u_i) & \Leftrightarrow & v_{i+1} \in V^0(u_{i+1}) \end{cases} \quad (3.4)$$

Because of their relative simplicity, studies of uniformly hyperbolic dynamical systems (also known as "ideal chaos") have provided a lot of insight into the properties of chaotic dynamical systems [19]. Although most real-life dynamical systems are not uniformly hyperbolic, they can be classified as *quasi-hyperbolic*: results obtained on hyperbolic systems can often be generalized to them [20]. This proof covers the convergence of LSS for uniform hyperbolic flows, nevertheless, numerical results have shown that the algorithm also works for non-ideal chaos [14].

#### 4. Neutral subspace and non-uniform discretization of trajectories.

One should bear in mind that the dynamical system is continuous which means that a solution  $\{u(t), t \in \mathbb{R}^+\}$  to equation (1.1) forms a continuous trajectory in phase space. Thus, the sequence  $\{u_i^{\{h, \infty\}}, i \in \mathbb{N}\}$  is no more than a sequence of sample points on the continuous trajectory, the time stepsize between two consecutive points being  $h$ . The *neutral subspace*  $V^0(u_i)$  which is unidimensional, is constituted by the line tangent to the continuous trajectory at the sampling point  $u_i$ .

Consequently, a perturbation in the direction of the *neutral subspace* around the point  $u_i$  can be interpreted as a time shift. This means that for  $i \in \mathbb{N}^*$  and  $\epsilon$  infinitesimal, if  $u(t_i) = u_i$  :

$$u(t_i + \epsilon) = u_i + \epsilon f(u_i, s) \quad (4.1)$$

and

$$u(t_i + \epsilon) = \varphi_s(u_{i-1}, h + \epsilon) = \varphi_s(u_{i-1} + \epsilon f(u_{i-1}, s), h) \quad (4.2)$$

which implies :

$$\varphi_s(u_{i-1} + \epsilon f(u_{i-1}, s), h) = u_i + \epsilon f(u_i, s) \quad (4.3)$$

i.e.

$$(D\varphi_s(u_{i-1}, t))f(u_{i-1}, s) = f(u_i, s) \quad (4.4)$$

Since  $\|f(u_i, s)\| \leq \sup_{s \in S} (\|f(\cdot, s)\|_\infty)^\Lambda < \infty$  for all  $i \in \mathbb{N}^*$ , then (4.4) implies that a small perturbation in the direction of the *neutral subspace* remains bounded under

the action of forward or backward iterations. On the contrary, a small perturbation in the *stable* or *unstable subspace* gets amplified exponentially under the action of backward or forward iterations respectively.

The second point to be discussed in this section is the fact that the discretization of a continuous trajectory doesn't have to respect a uniform stepsizing. A better way to discretize a trajectory would be the following one :  $\{(u_i, \tau_i), i \in \mathbb{N}\}$  is a sampling of a continuous trajectory if  $u_i = \varphi_s(u_{i-1}, h\tau_i)$  where  $\tau_i \in \mathbb{R}$ . In order to have :

$$\left| \langle J \rangle(s) - \lim_{T \rightarrow +\infty} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{\tau_i J(u_i, s)}{\sum_j \tau_j^s} \right| = O(h) \quad (4.5)$$

we only need  $\sup(h\tau_i) \rightarrow 0$  as  $h \rightarrow 0$  which actually happens if, for instance, the time dilatation factors  $\tau_i$  are bounded ( $\sum_j$  means  $\sum_{j=1}^{\lfloor \frac{T}{h} \rfloor}$  in the above expression). From now on, a discretization of a trajectory is a sequence of couples  $(u_i, \tau_i)$ .

**5. Structural stability and the shadowing direction.** In this section, we will prove a variant of the shadowing lemma for the purpose of defining the shadowing direction and prove its existence and uniqueness.

The hyperbolic structure ensures the structural stability of the attractor  $\Lambda$  under perturbation in  $s$  [21], [22]. Without loss of generality, we will assume that  $s = 0$  and choose a sequence  $\{u_i^0, i \in \mathbb{Z}\}$  such that :

$$u_{i+1}^0 = \varphi_0(u_i^0, h)$$

In this case, the superscript in  $u_i^0$  is the value of the parameter  $s$ .  $h$  and  $T = \infty$  are fixed so they do not appear in the notation.

**THEOREM 5.1 (Shadowing trajectory).** *If the system is uniformly hyperbolic and  $\varphi_s$  continuously differentiable with respect to  $s$  and  $u$ , then for all sequence  $\{u_i^0, i \in \mathbb{Z}\} \subset \Lambda$  satisfying  $u_i^0 = \varphi_0(u_{i-1}^0, h)$ , there is a  $M > 0$  such that for all  $|s| < M$  there is a sequence  $\{(u_i^s, \tau_i^s), i \in \mathbb{Z}\} \subset \mathbb{R}^m$  satisfying  $\|u_i^s - u_i^0\| < M$ ,  $\|\tau_i^s\| < M$  and  $u_i^s = \varphi_s(u_{i-1}^s, h\tau_i^s)$  for all  $i \in \mathbb{Z}$ . Furthermore,  $u_i^s$  and  $\tau_i^s$  are  $i$ -uniformly continuously differentiable to  $s$ .*

The  $i$ -uniform continuous differentiability of  $u_i^s$  means that for all  $s \in (-M, M)$  and  $\epsilon > 0$  there is a  $\delta > 0$  such that if  $|s - s'| < \delta$  then  $\|\frac{du_i^s}{ds}(s) - \frac{du_i^s}{ds}(s')\| < \epsilon$  and  $|\frac{d\tau_i^s}{ds}(s) - \frac{d\tau_i^s}{ds}(s')| < \epsilon$  for all  $i$ .

To prepare for the proof, let  $\mathbf{B}$  be the space of bounded sequences in  $\mathbb{R}^m$  and  $V_i$  the hyperplane of  $\mathbb{R}^m$  defined by  $V_i = V^+(u_i^0) \oplus V^-(u_i^0)$ . We introduce  $\mathbf{V}$  as the space of bounded sequences  $\{v_i, i \in \mathbb{Z}\}$  such that  $v_i \in V_i$  for all  $i \in \mathbb{Z}$  ( $v_i$  has no components in the *neutral subspace*). In other words :

$$\mathbf{V} = \left( \prod_{i \in \mathbb{Z}} V_i \right) \cap \mathbf{B}$$

Finally, by considering the space  $\mathbf{T}$  of bounded sequences of  $\mathbb{R}$ , we denote  $\mathbf{A}$  the product of  $\mathbf{V}$  by  $\mathbf{T}$  :

$$\mathbf{A} = \mathbf{V} \times \mathbf{T}$$

We then introduce the notation  $(\mathbf{v}, \boldsymbol{\tau}) = \{(v_i, \tau_i), i \in \mathbb{Z}\} \in \mathbf{A}$  where  $\mathbf{v} \in \mathbf{V}$ ,  $\boldsymbol{\tau} \in \mathbf{T}$  and define the norm :

$$\|(\mathbf{v}, \boldsymbol{\tau})\|_{\mathbf{A}} = \sup_{i \in \mathbb{Z}}(\|v_i\|) + \sup_{i \in \mathbb{Z}}(|\tau_i|) = \|\mathbf{v}\|_{\infty} + \|\boldsymbol{\tau}\|_{\infty}$$

As defined above, the space  $\mathbf{A}$  is a Banach space.

We can now define the map  $F : \mathbf{A} \times \mathbb{R} \rightarrow \mathbf{B}$  as :

$$\forall(\mathbf{v}, \boldsymbol{\tau}) \in \mathbf{A}, \forall s \in \mathbb{R}, \quad F((\mathbf{v}, \boldsymbol{\tau}), s) = \{u_i^0 + v_i - \varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1}), i \in \mathbb{Z}\}$$

For a given  $s$ ,  $F((\mathbf{v}, \boldsymbol{\tau}), s) = \mathbf{0}$  if and only if  $\{(u_i^0 + v_i), i \in \mathbb{Z}\}$  samples, with timesteps  $\{h\tau_i, i \in \mathbb{Z}\}$ , a continuous trajectory satisfying (1.1). We use the implicit function theorem to complete the proof, which requires  $F$  to be differentiable with respect to  $(\mathbf{v}, \boldsymbol{\tau})$  and its derivative to be non-singular at  $\mathbf{v} = \mathbf{0}$ ,  $\boldsymbol{\tau} = \mathbf{1}$  and  $s = 0$ .

LEMMA 5.2. *Under the conditions of theorem 5.1,  $F$  has Fréchet derivative at all  $(\mathbf{v}, \boldsymbol{\tau}) \in \mathbf{A}$  and  $|s| < M$ :*

$$(DF((\mathbf{v}, \boldsymbol{\tau}), s))(\mathbf{w}, \boldsymbol{\epsilon}) = \{w_i - (D\varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1}))(w_{i-1}) \quad (5.1)$$

$$- h\epsilon_{i-1}\partial_h\varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1}), i \in \mathbb{Z}\} \quad (5.2)$$

where  $(\mathbf{w}, \boldsymbol{\epsilon}) \in \mathbf{A}$

*Proof.* We have:

$$\frac{F((\mathbf{v} + \mathbf{w}, \boldsymbol{\tau} + \boldsymbol{\epsilon}), s) - F((\mathbf{v}, \boldsymbol{\tau}), s)}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} = \left\{ \frac{w_i}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} \quad (5.3)$$

$$- \frac{\varphi_s(u_{i-1}^0 + v_{i-1} + w_{i-1}, h(\tau_{i-1} + \epsilon_{i-1})) - \varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1})}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} \right\} \quad (5.4)$$

$$\xrightarrow{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}} \rightarrow 0} \left\{ \frac{w_i}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} - \left( \frac{D\varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1})(w_{i-1})}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} \right) \quad (5.5)$$

$$- \frac{h\epsilon_{i-1}\partial_h\varphi_s(u_{i-1}^0 + v_{i-1}, h\tau_{i-1})}{\|(\mathbf{w}, \boldsymbol{\epsilon})\|_{\mathbf{A}}} \right\} \quad (5.6)$$

in the  $\mathbf{A}$  norm thanks to the uniform continuity of  $D\varphi$  and  $\partial_h\varphi$  on the compact set  $\Lambda$ . Now, we only need to prove that the linear map we obtained is bounded. Since  $D\varphi$  and  $\partial_h\varphi$  are continuous they are uniformly bounded in the compact set  $\Lambda$ . Thus, the norm of the linear map is less than  $(1 + \|D\varphi\|_{\infty}^{\Lambda} + \|\partial_h\varphi\|_{\infty}^{\Lambda})$ .  $\square$

LEMMA 5.3. *Under conditions of theorem 5.1, the Fréchet derivative of  $F$  at  $(\mathbf{v}, \boldsymbol{\tau}) = (\mathbf{0}, \mathbf{1})$  and  $s = 0$  is a bijection.*

*Proof.* The Fréchet derivative of  $F$  at  $(\mathbf{v}, \boldsymbol{\tau}) = (\mathbf{0}, \mathbf{1})$  and  $s = 0$  in the direction  $(\mathbf{w}, \boldsymbol{\epsilon})$  is :

$$(DF((\mathbf{0}, \mathbf{1}), 0))(\mathbf{w}, \boldsymbol{\epsilon}) = \{w_i - (D\varphi_0(u_{i-1}^0, h))w_{i-1} - h\epsilon_{i-1}\partial_h\varphi_0(u_{i-1}^0, h), i \in \mathbb{Z}\}$$

To prove its bijectivity, we only need to show that for any  $\mathbf{r} = \{r_i\} \in \mathbf{B}$  there is a unique  $(\mathbf{w}, \boldsymbol{\epsilon}) \in \mathbf{A}$  such that  $(DF((\mathbf{0}, \mathbf{1}), 0))(\mathbf{w}, \boldsymbol{\epsilon}) = \mathbf{r}$

In this case, we can find an analytical expression for the pre-image of  $\mathbf{r}$ . Let  $(\mathbf{w}, \boldsymbol{\epsilon})$  be defined as :

$$(w_i, \epsilon_i) = \left( \sum_{j=0}^{\infty} (D\varphi_0(u_{i-j}^0, jh))r_{i-j}^- - \sum_{j=1}^{\infty} (D\varphi_0(u_{i+j}^0, -jh))r_{i+j}^+, \quad -\frac{1}{h} \frac{\langle r_{i+1}^0; \partial_h\varphi_0(u_i^0, h) \rangle}{\langle \partial_h\varphi_0(u_i^0, h); \partial_h\varphi_0(u_i^0, h) \rangle} \right) \quad (5.7) \quad \blacksquare$$

We can verify that  $w_i - (D\varphi_0(u_{i-1}^0, h))(w_{i-1}) - h\epsilon_{i-1}\partial_h\varphi_0(u_{i-1}^0, h) = r_i$  for all  $i$ .<sup>1</sup>

We still have to ensure that  $(\mathbf{w}, \epsilon)$  belongs to  $\mathbf{A}$ . Based on (3.4), we notice that the  $w_i$  we've just defined belongs to  $V_i = V^+(u_i^0) \oplus V^-(u_i^0)$ . Then, thanks to (3.1), we can write  $r_i = r_i^+ + r_i^- + r_i^0$  where  $r_i^+ \in V^+(u_i^0)$ ,  $r_i^- \in V^-(u_i^0)$  and  $r_i^0 \in V^0(u_i^0)$ . Since  $V^+(u)$ ,  $V^-(u)$  and  $V^0(u)$  are continuous with respect to  $u$  and  $\Lambda$  is compact:

$$\max(\|r_i^+\|, \|r_i^-\|, \|r_i^0\|) \leq \frac{\|r_i\|}{\beta} \leq \frac{\|\mathbf{r}\|_{\mathbf{B}}}{\beta} \quad \text{for all } i \quad (5.8)$$

where  $\beta > 0$ .

Consequently, for all  $i$ :

$$\|w_i\| \leq \sum_{j=0}^{\infty} \|(D\varphi_0(u_{i-j}^0, jh))r_{i-j}^-\| + \sum_{j=1}^{\infty} \|(D\varphi_0(u_{i+j}^0, -jh))r_{i+j}^+\| \quad (5.9)$$

$$\leq \sum_{j=0}^{\infty} C\lambda^{jh}\|r_{i-j}^-\| + \sum_{j=1}^{\infty} C\lambda^{jh}\|r_{i+j}^+\| \quad (5.10)$$

$$\leq \frac{2C}{1-\lambda^h} \frac{\|\mathbf{r}\|_{\mathbf{B}}}{\beta} \quad (5.11)$$

because  $V^+$  and  $V^-$  are invariant under  $D\varphi_0$  (property (3.4)). Thus,  $w_i$  is uniformly bounded and  $\mathbf{w} \in \mathbf{V}$ . In the same way, we show that for all  $i$ :

$$\epsilon_i \leq \frac{1}{h} \frac{\|r_{i+1}^0\| \|\partial_h\varphi_0(u_i^0, h)\|}{\|\partial_h\varphi_0(u_i^0, h)\|^2} \leq \frac{\|r_{i+1}^0\|}{h\|\partial\varphi_0(u_i^0, h)\|} \quad (5.12)$$

$$\leq \frac{\|\mathbf{r}\|_{\mathbf{B}}}{h\beta m} \quad (5.13)$$

where  $m = \inf_{u \in \Lambda} \{f(u, 0)\} > 0$ . Consequently,  $\epsilon_i$  is uniformly bounded which leads to  $\epsilon \in \mathbf{T}$  and  $(\mathbf{w}, \epsilon) \in \mathbf{A}$ .

Because of linearity, uniqueness of  $(\mathbf{w}, \epsilon)$  such that  $(DF((\mathbf{0}, \mathbf{1}), 0))(\mathbf{w}, \epsilon) = \mathbf{r}$  only needs to be proved for  $\mathbf{r} = \mathbf{0}$ . Since  $\mathbf{R}^m = V^+(u_i^0) \oplus V^-(u_i^0) \oplus V^0(u_i^0)$ ,  $r_i = 0$  is equivalent to  $r_i^+ = r_i^- = r_i^0 = 0$ . Thanks to property (3.4), by splitting  $w_i = w_i^+ + w_i^-$  and knowing that  $h\epsilon_{i-1}\partial_h\varphi_0(u_{i-1}^0, h) \in V^0(u_i^0)$ , we have:

$$0 = r_i^+ + r_i^- = (w_i^+ - (D\varphi_h(u_{i-1}^0, 0))w_{i-1}^+) + (w_i^- - (D\varphi_h(u_{i-1}^0, 0))w_{i-1}^-) \quad (5.14)$$

where the two parentheses are in  $V^+(u_i^0)$  and  $V^-(u_i^0)$  respectively. Again knowing that  $\mathbf{R}^m = V^+(u_i^0) \oplus V^-(u_i^0) \oplus V^0(u_i^0)$ , both parentheses should be equal to zero. This is true for all  $i$ , so we obtain :

$$w_i^+ = (D\varphi_0(u_{i-1}^0, h))w_{i-1}^+ = \dots = (D\varphi_0^{(i-j)}(u_j^0, h))w_j^+ \quad (5.15)$$

$$w_i^- = (D\varphi_0(u_{i-1}^0, h))w_{i-1}^- = \dots = (D\varphi_0^{(i-j)}(u_j^0, h))w_j^- \quad (5.16)$$

for all  $j < i$ . Based on the properties of uniform hyperbolicity,  $\|w_j^+\| \leq C\lambda^{h(i-j)}\|w_i^+\|$  and  $\|w_j^-\| \leq C\lambda^{h(i-j)}\|w_i^-\|$ . If for some  $j$  we have  $w_j^+ \neq 0$ , then:

$$\frac{\|w_i\|}{\beta} \geq \|w_i^+\| \geq \frac{\lambda^{h(j-i)}}{C} \|w_j^+\| \quad \text{for all } i > j \quad (5.17)$$

---

<sup>1</sup>Based on the fact that  $D\varphi_0(u_{i-1}^0, h)D\varphi_0(u_{i-j-1}^0, jh) = D\varphi_0(u_{i-j-1}^0, (j+1)h)$  and  $D\varphi_0(u_{i-1}^0, h)D\varphi_0(u_{i-1+j}^0, -jh) = D\varphi_0(u_{i-1+j}^0, (-j+1)h)$



which means that  $\{w_i, i \in \mathbf{Z}\}$  is unbounded ( $0 < \lambda < 1$ ). Similarly, if  $v_i^- \neq 0$  for some  $i$  then:

$$\frac{\|w_j\|}{\beta} \geq \|w_j^+\| \geq \frac{\lambda^{h(j-i)}}{C} \|w_i^+\| \quad \text{for all } j < i \quad (5.18)$$

and  $\{w_i, i \in \mathbf{Z}\}$  is also unbounded. Consequently, for  $\{w_i\}$  to be bounded we must have  $w_i = w_i^+ + w_i^- = 0$  for all  $i$ .

On the other hand, showing that  $\epsilon_i = 0$  is trivial:

$$0 = r_i^0 = -h\epsilon_{i-1}\partial_h\varphi_0(u_i^0, h) \quad (5.19)$$

Since  $\|\partial_h\varphi_0(u_i^0, h)\| \geq m > 0$  then  $\epsilon_{i-1} = 0$ . This is true for all  $i$ , which means that  $\epsilon = 0$ . This proves the uniqueness of  $(\mathbf{w}, \epsilon)$  for  $\mathbf{r} = 0$ .  $\square$

*Proof. (of theorem 5.1)* Since  $F((\mathbf{0}, 1), 0) = \{u_i^0 - \varphi_0(u_{i-1}^0, h)\} = \mathbf{0}$ ,  $(\mathbf{0}, 1)$  is a zero point of  $F$  at  $s = 0$ . Based on this information and on the two previous lemmas, the implicit function theorem states that there exist  $M > 0$  such that for all  $|s| < M$  there is a unique  $(\mathbf{v}^s, \boldsymbol{\tau}^s)$  satisfying  $\|(\mathbf{v}^s, \boldsymbol{\tau}^s)\|_{\mathbf{A}} < M$  and  $F((\mathbf{v}^s, \boldsymbol{\tau}^s), s) = \mathbf{0}$ . Furthermore, this  $(\mathbf{v}^s, \boldsymbol{\tau}^s)$  is continuously differentiable to  $s$ , i.e.,  $\frac{d(\mathbf{v}^s, \boldsymbol{\tau}^s)}{ds} \in \mathbf{A}$  is continuous with respect to  $s$  in the  $\mathbf{A}$  norm. By the definition of derivatives (in  $\mathbf{A}$ ),  $\frac{d(\mathbf{v}^s, \boldsymbol{\tau}^s)}{ds} = \{(\frac{dv_i^s}{ds}, \frac{d\tau_i^s}{ds})\}$ . Continuity of  $\frac{d(\mathbf{v}^s, \boldsymbol{\tau}^s)}{ds}$  in  $\mathbf{A}$  then implies that  $\frac{dv_i^s}{ds}$  and  $\frac{d\tau_i^s}{ds}$  are  $i$ -uniformly continuous with respect to  $s$ . By defining:

$$\{(u_i^s, \tau_i^s), i \in \mathbf{Z}\} = \{(u_i^0 + v_i^s, \tau_i^s), i \in \mathbf{Z}\} \quad (5.20)$$

we finally obtain the results of theorem (5.1).  $\square$

If we return to the expanded notation of  $u_i^s$ , this theorem states that for a discretization  $\{(u_i^{0\{h,\infty\}}, 1)\}$  of a continuous trajectory satisfying (1.1) for  $s = 0$ , there is a series  $\{(u_i^{s\{h,\infty\}}, \tau_i^{s\{h,\infty\}})\}$  also being a discretization of a continuous trajectory at nearby values of  $s$ . In addition,  $\{(u_i^{s\{h,\infty\}}, \tau_i^{s\{h,\infty\}})\}$  *shadows*  $\{(u_i^{0\{h,\infty\}}, 1)\}$  meaning that  $\{(u_i^{s\{h,\infty\}}, \tau_i^{s\{h,\infty\}})\}$  is close to  $\{(u_i^{0\{h,\infty\}}, 1)\}$  when  $s$  is close to 0. Also,  $\{(\frac{du_i^{s\{h,\infty\}}}{ds}, \frac{d\tau_i^{s\{h,\infty\}}}{ds})\}$  exists and is  $i$ -uniformly bounded.

The *shadowing direction*  $(v_i^{\{h,\infty\}}, \eta_i^{\{h,\infty\}})$  is defined as the uniformly bounded series :

$$\{(v_i^{\{h,\infty\}}, \eta_i^{\{h,\infty\}})\} := \left\{ \left( \frac{du_i^{s\{h,\infty\}}}{ds} \Big|_{s=0}, \frac{d\tau_i^{s\{h,\infty\}}}{ds} \Big|_{s=0} \right) \right\} \in \mathbf{A} \quad (5.21)$$

In addition, we can find 2 constants  $\|\mathbf{v}^{\{\infty\}}\|$  and  $\|\boldsymbol{\eta}^{\{\infty\}}\|$  independant of  $h$  such that for all  $i$ :

$$v_i^{\{h,\infty\}} \leq \|\mathbf{v}^{\{\infty\}}\| \quad \text{and} \quad \eta_i^{\{h,\infty\}} \leq \|\boldsymbol{\eta}^{\{\infty\}}\| \quad (5.22)$$

Please refer to appendix A to see how these constants are built.

**6. A simpler result.** In this section, we prove an easier version of Theorem LSS in which we replace the solution  $\{(v_i^{\{h,T\}}, \eta_i^{\{h,T\}}), i = 1, \dots, [\frac{T}{h}]\}$  to the constrained least squares problem (2.3) by the shadowing direction we found earlier  $\{(v_i^{\{\infty\}}, \eta_i^{\{\infty\}}), i = 1, \dots, [\frac{T}{h}]\}$  which can be written  $\{(v_i^{\{h,\infty\}}, \eta_i^{\{h,\infty\}}), i = 1, \dots, [\frac{T}{h}]\}$  if  $h$  is to be displayed explicitly.

**THEOREM 6.1.** *If uniform hyperbolicity holds and  $\varphi_s(\cdot, h)$  is continuously differentiable for all  $h$ , then for all continuously differentiable function  $J : \mathbf{R}^m \times \mathbf{R} \rightarrow \mathbf{R}$  whose infinite time average :*

$$\langle J \rangle(s) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T J(u(t), s) dt \quad \text{where} \quad \frac{du}{dt} = f(u, s) \quad \text{and} \quad u(0) = u_0 \quad (6.1)$$

is independant of the initial state  $u_0$ , let  $\{(v_i^{\{h,\infty\}}, \eta_i^{\{h,\infty\}}), i = 1, \dots, [\frac{T}{h}]\}$  be the sequence of shadowing direction, then:

$$\frac{d\langle J \rangle}{ds} = \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{h}{T} \sum_{i=1}^{[\frac{T}{h}]} \left[ (DJ(u_i, 0))v_i^{\{h,\infty\}} + (\partial_s J(u_i, 0)) + (\eta_i^{\{h,\infty\}}(J(u_i, 0) - \langle J \rangle(0))) \right] \quad (6.2)$$

$$= \lim_{T \rightarrow \infty} \lim_{h \rightarrow 0} \frac{h}{T} \sum_{i=1}^{[\frac{T}{h}]} \left[ (DJ(u_i, 0))v_i^{\{h,\infty\}} + (\partial_s J(u_i, 0)) + (\eta_i^{\{h,\infty\}}(J(u_i, 0) - \langle J \rangle(0))) \right] \quad (6.3) \quad \blacksquare$$

*Proof.* The proof is essentially an exchange of limits through uniform convergence. Since  $\langle J \rangle$  is independant of  $u_0$ , we set  $u_0 = u_0^{s\{h,\infty\}}$  as defined in the previous section and we know that  $u_{i+1}^{s\{h,\infty\}} = \varphi_s(u_i^{s\{h,\infty\}}, h\tau_i^{s\{h,\infty\}})$ . In order to simplify the notations,  $u_i^{s\{h,\infty\}}$  and  $\tau_i^{s\{h,\infty\}}$  will be expressed as  $u_i^s$  and  $\tau_i^s$  in what follows. The integral could be discretized into an infinite sommation that involves the time dilatation factors:

$$\langle J \rangle(s) = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T J(u(t), s) dt = \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \sum_{i=1}^{[\frac{T}{h}]} \frac{\tau_i^s J(u_i^s, s)}{\sum_j \tau_j^s} \quad (6.4)$$

$$= \lim_{T \rightarrow +\infty} \lim_{h \rightarrow 0} \sum_{i=1}^{[\frac{T}{h}]} \frac{\tau_i^s J(u_i^s, s)}{\sum_j \tau_j^s} \quad (6.5)$$

The two limits can be permuted at will but we will only keep one notation for the remaining of the proof. We can write:

$$\begin{aligned}
\left. \frac{d\langle J \rangle}{ds} \right|_{s=0} &= \lim_{s \rightarrow 0} \frac{\langle J \rangle(s) - \langle J \rangle(0)}{s} \\
&= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{1}{s} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} - \frac{h J(u_i^0, 0)}{T} \right) \\
&= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{1}{s} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} - \frac{h J(u_i^0, 0)}{h \sum_j \tau_j^s} + \frac{h J(u_i^0, 0)}{h \sum_j \tau_j^s} - \frac{h J(u_i^0, 0)}{T} \right) \\
&= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{1}{s} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{h(J(u_i^s, s) - J(u_i^0, 0) + (\tau_i^s - 1)J(u_i^s, s))}{h \sum_j \tau_j^s} \right. \\
&\quad \left. - \frac{(h \sum_j (\tau_j^s - 1)) \times h J(u_i^0, 0)}{T(h \sum_j \tau_j^s)} \right) \\
&= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{1}{s} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{h(J(u_i^s, s) - J(u_i^0, 0) + (\tau_i^s - 1)J(u_i^s, s))}{T + h \sum_j (\tau_j^s - 1)} \right. \\
&\quad \left. - \frac{(h \sum_j (\tau_j^s - 1)) \times h J(u_i^0, 0)}{T(T + h \sum_j (\tau_j^s - 1))} \right) \\
&= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{(J(u_i^s, s) - J(u_i^0, 0))}{s} + O(s) \right) \\
&\quad + \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h(\tau_i^s - 1)}{Ts} (J(u_i^s, s) - \frac{1}{T} \sum_j h J(u_i^0, 0)) + O(s) \right)
\end{aligned}$$

Let us eliminate  $\lim_{s \rightarrow 0}$  in the first term. We define :

$$\gamma_i^s = \frac{dJ(u_i^s, s)}{ds} = (DJ(u_i^s, s)) \frac{du_i^s}{ds} + \partial_s J(u_i^s, s) \quad (6.6)$$

Then, thanks to the mean value theorem, for all  $i$  there exist an  $\xi_i(s) \in [0, s]$  such that:

$$\frac{(J(u_i^s, s) - J(u_i^0, 0))}{s} = \gamma_i^{\xi_i(s)} \quad (6.7)$$

Consequently:

$$\lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \frac{(J(u_i^s, s) - J(u_i^0, 0))}{s} \right) = \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^{\xi_i(s)} \right) \quad (6.8)$$

We can choose a neighborhood of  $\Lambda \times \{0\}$  that contains  $(u_i^s, s)$  for all  $i$  (for  $s$  sufficiently small) and in which both  $(DJ(u, s))$  and  $\partial_s J(u, s)$  are uniformly continuous. Since the  $\frac{du_i^s}{ds}$  are  $i$ -uniformly continuous and bounded, for all  $\epsilon > 0$  there exists  $M > 0$  such that for all  $|\xi| < M$ :

$$\|\gamma_i^\xi - \gamma_i^0\| < \epsilon \quad \forall i$$

Thus, for all  $|s| < M$ ,  $|\xi_i(s)| \leq |s| < M$  for all  $i$ , therefore for all  $h, T$  ( $h \ll T$ ) :

$$\left\| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^{\xi_i(s)} - \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^0 \right\| \leq \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|\gamma_i^{\xi_i(s)} - \gamma_i^0\| < \epsilon \quad (6.9)$$

Hence,

$$\left\| \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^{\xi_i(s)} \right) - \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^0 \right) \right\| \leq \epsilon \quad (6.10)$$

Finally,

$$\lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^{\xi_i(s)} \right) = \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \gamma_i^0 \right) \quad (6.11)$$

which grants us the desired result for the first term via the definition of  $\gamma_i^0$ .

For the second term,  $J$  is continuously differentiable thus continuous and the  $(u_i^s, \tau_i^s)$  are  $i$ -uniformly continuously differentiable and bounded. Based on that, for  $s$  sufficiently small, we can find a compact neighborhood of  $\Lambda \times \{0\}$  that contains  $(u_i^s, s)$  for all  $i \in \mathbf{Z}$  and in which  $J(u, s)$  will be uniformly continuous. Consequently, the sequence  $\left\{ \frac{h(\tau_i^s - 1)}{Ts} J(u_i^s, s), i \in \mathbf{Z} \right\}$  which can be written  $\left\{ \frac{h(\tau_i^s - \tau_i^0)}{Ts} J(u_i^s, s), i \in \mathbf{Z} \right\}$  converges uniformly to  $\left\{ \frac{h\eta_i^{\{h, \infty\}}}{T} J(u_i^0, s), i \in \mathbf{Z} \right\}$  when  $s$  goes to 0. Because the term  $\frac{1}{T} \sum_j hJ(u_i^0, 0)$  does not depend on  $s$  at all, we finally have:

$$\lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h(\tau_i^s - 1)}{Ts} (J(u_i^s, s) - \frac{1}{T} \sum_j hJ(u_i^0, 0)) \right) = \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{h\eta_i^{\{h, \infty\}}}{T} (J(u_i^0, 0) - \langle J \rangle(0)) \right) \quad \blacksquare$$

which concludes the proof.

□

**7. Computational approximation of the shadowing direction.** The main task of this section is to provide a bound for :

$$e_i^{\{h, T\}} = v_i^{\{h, T\}} - v_i^{\{h, \infty\}} \quad (7.1)$$

$$\epsilon_i^{\{h, T\}} = \eta_i^{\{h, T\}} - \eta_i^{\{h, \infty\}} \quad (7.2)$$

for  $i = 1, \dots, \frac{T}{h}$  where the  $v_i^{\{h, T\}}, \eta_i^{\{h, T\}}$  are the solution to the least squares problem:

$$\min \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h, T\}}\|^2 + \alpha(\eta_i^{\{h, T\}})^2) \quad (7.3)$$

$$\text{s.t. } v_{i+1}^{\{h, T\}} = (D\varphi_s(u_i, h))v_i^{\{h, T\}} + (\partial_s \varphi_s(u_i, h)) + h\eta_i^{\{h, T\}} \partial_h \varphi_s(u_i, h), \quad (7.4)$$

The shadowing lemma guarantees the existence of a shadowing trajectory, but provides no clear way to compute  $\{(v_i^{\{h, \infty\}}, \eta_i^{\{h, \infty\}})\}$ . This section suggests that the solution to the least squares problem gives a useful approximation of the shadowing

trajectory allowing us to compute  $\frac{d\langle J \rangle}{ds}$ . Without loss of generality, we consider that  $s = 0$  in (7.4). By definition, the shadowing trajectory satisfies:

$$u_{i+1}^{s\{h,\infty\}} = \varphi_s(u_i^{s\{h,\infty\}}, h\tau_i^{s\{h,\infty\}}) \quad (7.5)$$

After taking the derivative to  $s$  on both sides for  $s = 0$ , we obtain:

$$v_{i+1}^{\{h,\infty\}} = (D\varphi_0(u_i, h))v_i^{\{h,\infty\}} + \frac{\partial \varphi_0}{\partial s}(u_i, h) + h\eta_i^{\{h,\infty\}} \partial_h \varphi_0(u_i, h) \quad (7.6)$$

Thus, the shadowing direction satisfies the constraint in equation (7.3) and:

$$\sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h,T\}}\|^2 + \alpha(\eta_i^{\{h,T\}})^2) \leq \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h,\infty\}}\|^2 + \alpha(\eta_i^{\{h,\infty\}})^2) \leq \frac{T}{h} (\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\boldsymbol{\eta}^{\{\infty\}}\|^2) \quad (7.7) \quad \blacksquare$$

Consequently :

$$\max_i \|\eta_i^{\{h,T\}}\| \leq \sqrt{\frac{T}{h}} \left( \frac{\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2}{\alpha} + \|\boldsymbol{\eta}^{\{\infty\}}\|^2 \right)^{\frac{1}{2}} \quad (7.8)$$

since  $\epsilon_i^{\{h,T\}} = \eta_i^{\{h,T\}} - \eta_i^{\{h,\infty\}}$  :

$$\max_i \|\epsilon_i^{\{h,T\}}\| \leq \sqrt{\frac{T}{h}} \left( \frac{\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2}{\alpha} + \|\boldsymbol{\eta}^{\{\infty\}}\|^2 \right)^{\frac{1}{2}} + \|\boldsymbol{\eta}^{\{\infty\}}\| \quad (7.9)$$

We can get a similar bound for  $\max_i \|e_i^{\{h,T\}0}\|$  :

$$\max_i \|e_i^{\{h,T\}0}\| \leq \sqrt{\frac{T}{h}} \frac{(\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\boldsymbol{\eta}^{\{\infty\}}\|^2)^{\frac{1}{2}}}{\gamma} + \frac{\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}}{\gamma} \quad (7.10)$$

Concerning  $\max_i \|e_i^{\{h,T\}+}\|$  and  $\max_i \|e_i^{\{h,T\}-}\|$ , combining the constraint equation (7.4) as well as (7.6) we obtain :

$$\begin{cases} e_{i+1}^{\{h,T\}+} = D\varphi_0(u_i, h)e_i^{\{h,T\}+} \\ e_{i+1}^{\{h,T\}-} = D\varphi_0(u_i, h)e_i^{\{h,T\}-} \end{cases}$$

Thus, since  $e_i^{\{h,T\}+} \in V^+(u_i)$ ,  $e_i^{\{h,T\}-} \in V^-(u_i)$  and based on the following definition of the *unstable* and *stable subspaces* :

$$V^+(u) := \{v \in \mathbb{R}^m / \|(D\varphi_s(u, t))v\| \leq C\lambda^{-t}\|v\|, \forall t < 0\} \quad (7.11)$$

$$V^-(u) := \{v \in \mathbb{R}^m / \|(D\varphi_s(u, t))v\| \leq C\lambda^t\|v\|, \forall t > 0\} \quad (7.12)$$

we obtain :

$$\max_i \|e_i^{\{h,T\}+}\| \leq \max(1, C)\|e_{\lfloor \frac{T}{h} \rfloor}^{\{h,T\}+}\| \quad (7.13)$$

$$\max_i \|e_i^{\{h,T\}-}\| \leq \max(1, C)\|e_1^{\{h,T\}-}\| \quad (7.14)$$

In addition to that,  $\varphi_0(u, 0) = u$  which means that  $D\varphi_0(u, 0) = I$  where  $I$  is the identity operator in  $\mathbb{R}^m$ . For  $u \in \Lambda$ ,  $h \in H$  and  $s \in S$  where  $H$  and  $S$  are compact sets containing  $h = 0$  and  $s = 0$ , we can find a positive constant  $K$  such that:

$$\|\partial_h D\varphi_0(u, h)\| \leq K \quad (7.15)$$

since  $\varphi(\cdot, \cdot)$  is  $C^2$ .  $\partial_h D\varphi_0(u, h)$  belongs to a space of finite dimension so all norms are equivalent and any choice of norm  $\|\cdot\|$  is valid. For a sufficiently small  $h$  such that  $hK < 1$ , we get :

$$\|(D\varphi_0(u, h))v\| \geq (1 - hK)\|v\| \quad (7.16)$$

for all  $v \in \mathbb{R}^m$ . Consequently :

$$\sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|e_i^{\{h, T\}^-}\|^2 \geq \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (1 - hK)^{2(i-1)} \|e_1^{\{h, T\}^-}\|^2 \quad (7.17)$$

$$\geq \|e_1^{\{h, T\}^-}\|^2 \frac{1 - (1 - hK)^{2\lfloor \frac{T}{h} \rfloor}}{h(2K - hK^2)} \quad (7.18)$$

which means that:

$$\|e_1^{\{h, T\}^-}\|^2 \leq \frac{h(2K - hK^2)}{1 - (1 - hK)^{2\lfloor \frac{T}{h} \rfloor}} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|e_i^{\{h, T\}^-}\|^2 \quad (7.19)$$

$$\leq \frac{h(2K - hK^2)}{1 - e^{-2TK}} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|e_i^{\{h, T\}^-}\|^2 \leq \frac{2hK}{1 - e^{-2TK}} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|e_i^{\{h, T\}^-}\|^2 \quad (7.20)$$

For  $T$  sufficiently big,  $e^{-2KT} < \frac{1}{2}$  which means that:

$$\|e_1^{\{h, T\}^-}\|^2 \leq 4hK \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \|e_i^{\{h, T\}^-}\|^2 \quad (7.21)$$

Since  $e_i^{\{h, T\}^-} = v_i^{\{h, T\}^-} - v_i^{\{h, \infty\}^-}$  then:

$$\|e_i^{\{h, T\}^-}\|^2 \leq 2(\|v_i^{\{h, T\}^-}\|^2 + \|v_i^{\{h, \infty\}^-}\|^2) \leq \frac{2}{\gamma}(\|v_i^{\{h, T\}}\|^2 + \|v_i^{\{h, \infty\}}\|^2) \quad (7.22)$$

leading to:

$$\|e_1^{\{h, T\}^-}\|^2 \leq \frac{8hK}{\gamma} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h, T\}}\|^2 + \|v_i^{\{h, \infty\}}\|^2) \quad (7.23)$$

$$\leq \frac{8hK}{\gamma} \times \frac{T}{h} (2\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\boldsymbol{\eta}^{\{\infty\}}\|^2) \quad (7.24)$$

Finally :

$$\max_i \|e_i^{\{h, T\}^-}\| \leq \sqrt{T} \times \max(1, C) \times \left( \frac{8K}{\gamma} \times (2\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\boldsymbol{\eta}^{\{\infty\}}\|^2) \right)^{\frac{1}{2}} \quad (7.25)$$

$$\leq E\sqrt{T} \quad (7.26)$$

where  $E = \max(1, C) \times \left( \frac{8K}{\gamma} \times (2\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\boldsymbol{\eta}^{\{\infty\}}\|^2) \right)^{\frac{1}{2}}$  is a constant that doesn't depend on  $h$  nor on  $T$ . In the same way we obtain :

$$\max_i \|e_i^{\{h,T\}+}\| \leq E\sqrt{T} \quad (7.27)$$

Even though the bounds we found for  $\eta_i^{\{h,T\}}$ ,  $e_i^{\{h,T\}0}$ ,  $e_i^{\{h,T\}+}$  and  $e_i^{\{h,T\}-}$  may depend on  $T$  and/or  $h$ , we will see in the next section that they are strong enough to prove the convergence of the algorithm. Furthermore, experimental simulations have shown that  $e_i^{\{h,T\}}$  and  $\eta_i^{\{h,T\}}$  doesn't necessarily grow when  $h \rightarrow 0$ ,  $T \rightarrow \infty$  and stay bounded [18].

**8. Convergence of least squares shadowing.** In this section, we use the results obtained previously to prove our initial theorem :

**THEOREM 8.1 (THEOREM LSS).** *For a  $C^2$  map  $\varphi(\cdot, \cdot)$  and a  $C^1$  cost function  $J$ , the following limit exists and is equal to:*

$$\begin{aligned} \frac{d\langle J \rangle}{ds} &= \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))v_i^{\{h,T\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \\ &= \lim_{T \rightarrow \infty} \lim_{h \rightarrow 0} \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))v_i^{\{h,T\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \end{aligned}$$

*Proof.* Because  $J$  is  $C^1$  and  $\Lambda$  is compact,  $(DJ(u_i, 0))$  is uniformly bounded, i.e., there exists a constant  $A$  such that  $\|DJ(u_i, 0)\| < A$  for all  $i$ . Let  $e_i^{\{h,T\}}$  be defined as in the previous section, then:

$$\left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))v_i^{\{h,T\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.1)$$

$$- \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))v_i^{\{h,\infty\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h,\infty\}}(J(u_i, s) - \langle J \rangle(s))) \right] \quad (8.2)$$

$$= \left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))e_i^{\{h,T\}} + (\epsilon_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.3)$$

$$= \left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))(e_i^{\{h,T\}+} + e_i^{\{h,T\}-} + e_i^{\{h,T\}0}) + (\epsilon_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.4)$$

$$< \left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))(e_i^{\{h,T\}+} + e_i^{\{h,T\}-}) \right] \right| + \left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))e_i^{\{h,T\}0} + (\epsilon_i^{\{h,T\}}(J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.5)$$

For the first term:

$$\left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))(e_i^{\{h,T\}+} + e_i^{\{h,T\}-}) \right] \right| < \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \| (DJ(u_i, s)) \| \| e_i^{\{h,T\}+} \| \quad (8.6)$$

$$+ \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \| (DJ(u_i, s)) \| \| e_i^{\{h,T\}-} \| \quad (8.7)$$

$$\leq A \frac{h}{T} \left( \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} C \lambda^{h(\frac{T}{h}-i)} \| e_{\lfloor \frac{T}{h} \rfloor}^{\{h,T\}+} \| + \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} C \lambda^{h(i-1)} \| e_0^{\{h,T\}-} \| \right) \quad (8.8)$$

$$\leq \frac{h}{T} \frac{2AC}{(1-\lambda^h)} \times E\sqrt{T} \quad (8.9)$$

$$\leq \frac{1}{\sqrt{T}} \times \frac{2ACE}{\log(\frac{e}{\lambda})} \quad (8.10)$$

which goes to 0 when  $T$  increases. Thus, we notice that, in the stable and unstable subspaces, the difference  $e_i^{\{h,T\}}$  between the shadowing trajectory  $v_i^{\{h,\infty\}}$  and its approximation  $v_i^{\{h,T\}}$  decreases extremely fast so that the whole term  $\left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s))(e_i^{\{h,T\}+} + e_i^{\{h,T\}-}) \right] \right|$  tends to 0.

The situation is more complicated for the second term since there is no reason for  $e_i^{\{h,T\}0}$  and  $\epsilon_i^{\{h,T\}}$  to decrease when  $\frac{T}{h}$  increases. The cancellation of the second term is the result of the mutual cancellation of the elements in the summation as we shall see. Based on the sampling points  $\{(u_i^{s\{h,\infty\}}, \tau_i^{s\{h,\infty\}})\}$  for the continuous shadowing trajectory found in section 5, we consider the new set of sampling points  $\{(u_i'^s, \tau_i^{s\{h,\infty\}} + s\epsilon_i^{\{h,T\}})\}$  which satisfy the following relation :

$$\lim_{s \rightarrow 0} \frac{u_i'^s - u_i^s}{s} = e_i^{\{h,T\}} \quad (8.11)$$

for all  $i$ . We can notice that the new sampling points describe the same continuous trajectory as the old set of values. Assuming that  $\epsilon_i^{\{h,T\}}$  and  $e_i^{\{h,T\}}$  are bounded, we have for a sufficiently small  $s$ :

$$\langle J \rangle(s) = \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{h(\tau_i^s + s\epsilon_i^{\{h,T\}})J(u_i'^s, s)}{h \sum_j (\tau_j^s + s\epsilon_j^{\{h,T\}})} \quad (8.12)$$

We would obtain by following the same operations we did in section 6 (but upside



down this time) :

$$\lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) e_i^{\{h, T\}0} + (\epsilon_i^{\{h, T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right) \quad (8.13)$$

$$= \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( \frac{h}{T} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{(J(u_i^s, s) - J(u_i^s, s))}{s} \right) \quad (8.14)$$

$$+ \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( \frac{h \epsilon_i^{\{h, T\}}}{T} \left( J(u_i^s, s) - \frac{1}{T} \sum_j h J(u_i^s, s) \right) \right) \quad (8.15)$$

$$= \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( \frac{1}{s} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{h(\tau_i^s + s \epsilon_i^{\{h, T\}}) J(u_i^s, s)}{h \sum_j (\tau_j^s + s \epsilon_j^{\{h, T\}})} - \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} \right) + O(s) \quad (8.16)$$

$$= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( \frac{1}{s} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{h(\tau_i^s + s \epsilon_i^{\{h, T\}}) J(u_i^s, s)}{h \sum_j (\tau_j^s + s \epsilon_j^{\{h, T\}})} - \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} \right) + O(s) \quad (8.17)$$

$$= \lim_{s \rightarrow 0} \frac{\langle J \rangle(s) - \langle J \rangle(s)}{s} + O(s) = 0 \quad (8.18)$$

However, this is not necessarily true since  $e_i^{\{h, T\}0}$  and  $\epsilon_i^{\{h, T\}}$  may grow as  $\frac{T}{h}$  increases. Permuting the limits is much more delicate but is still possible. Please refer to appendix B for further details about how to permute the limits. The idea is to use the relation :

$$\sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h, T\}}\|^2 + \alpha(\eta_i^{\{h, T\}})^2) \leq \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h, \infty\}}\|^2 + \alpha(\eta_i^{\{h, \infty\}})^2) \leq \frac{T}{h} (\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \alpha\|\eta^{\{\infty\}}\|^2) \quad (8.19)$$

to show that most  $e_i^{\{h, T\}0}$ ,  $\epsilon_i^{\{h, T\}}$  remain bounded and that the contribution of the unbounded terms fades out.

In conclusion :

$$\left| \frac{d\langle J \rangle}{ds} - \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, T\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h, T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.20)$$

$$\leq \left| \frac{d\langle J \rangle}{ds} - \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, \infty\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h, \infty\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.21)$$

$$+ \left| \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, T\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h, T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right| \quad (8.22)$$

$$- \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i, s)) v_i^{\{h, \infty\}} + (\partial_s J(u_i, s)) + (\eta_i^{\{h, \infty\}} (J(u_i, s) - \langle J \rangle(s))) \right] \quad (8.23)$$

$$(8.24)$$

and we have showed that both terms go to 0 as  $T \rightarrow \infty$  and  $h \rightarrow 0$ . This concludes the proof.  $\square$

**9. Conclusion.** As we have shown through this paper, LSS gives us a good estimation for  $\frac{d\langle J \rangle}{ds}$  when the dynamical system is uniformly hyperbolic. After running a simulation for a given  $s$ , an arbitrary initial condition  $u_0$  and an uniform time stepsize of  $h$ , we obtain a sequence of reference sampling points<sup>2</sup>  $\{u_i^s, i = 1, \dots, \frac{T}{h}\}$ . If we had access to the *shadowing direction*, we would easily compute :

$$\frac{d\langle J \rangle}{ds} \approx \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \left[ (DJ(u_i^s, s))v_i^{\{h, \infty\}} + (\partial_s J(u_i^s, s)) + (\eta_i^{\{h, \infty\}}(J(u_i^s, s) - \langle J \rangle(s))) \right] \quad (9.1)$$

However, in real-life problems we usually do not have access to the *stable* and *unstable subspaces* around each  $u_i^s$  prohibiting the usage of the closed form expression of  $v_i^{\{h, \infty\}}$  and  $\eta_i^{\{h, \infty\}}$ . Thus, we have no other choice than computing an approximation of the *shadowing direction*. This approximation is given by the solution to the least squares problem:

$$\begin{aligned} & \min \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (\|v_i^{\{h, T\}}\|^2 + \alpha(\eta_i^{\{h, T\}})^2) \\ \text{s.t. } & v_{i+1}^{\{h, T\}} = (D\varphi_s(u_i^s, h))v_i^{\{h, T\}} + \partial_s \varphi_s(u_i^s, h) + h\eta_i^{\{h, T\}} \partial_h \varphi_s(u_i^s, h), \end{aligned} \quad (9.2)$$

After solving this quadratic optimization problem, we estimate  $\frac{d\langle J \rangle}{ds}$  using expression (9.1) again where the  $(v_i^{\{h, \infty\}}, \eta_i^{\{h, \infty\}})$  are replaced by  $(v_i^{\{h, T\}}, \eta_i^{\{h, T\}})$ . As we have seen previously, this estimation converges to the real value of  $\frac{d\langle J \rangle}{ds}$  when the time stepsize  $h$  is refined and the integration lapse  $T$  increases.

## REFERENCES

- [1] D. RUELLE, *Differentiation of SRB states*, Communications in Mathematical Physics, 187:227-241, 1997.
- [2] R. BUIZZA, *Chaos and weather prediction January 2000*, Analysis 12, 1-7, 1996.
- [3] C.S DAW, J.F. THOMAS, G.A. RICHARDS, L.L. NARAYANASWAMI, *Chaos in thermal pulse combustion*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 5(4), 662-670, 1995.
- [4] C.J. CHANG, R.T. LAHEY, *Analysis of chaotic instabilities in boiling systems*, Nuclear Engineering and Design 167.3, 307-334, 1997.
- [5] T. KLINGER, C. SCHRÖDER, D. BLOCK, F. GREINER, A. PIEL, G. BONHOMME, V. NAULIN, *Chaos control and taming of turbulence in plasma devices*, Physics of Plasmas, 8(5) 1961-1968, 2001.
- [6] C. FROESCHLÉ, *Modelling: An aim and a tool for the study of the chaotic behaviour of asteroidal and cometary orbits*, Predictability, Stability, and Chaos in N-Body Dynamical Systems, Springer, 125-155, 1991.
- [7] J. REUTHER, A. JAMESON, J. ALONSO, M. RIMLINGER, D. SAUNDERS, *Constrained multi-point aerodynamic shape optimization using an adjoint formulation and parallel computers*, Journal of aircraft(1), 51-60, 1999.
- [8] Q. WANG, *Uncertainty quantification for unsteady fluid flow using adjoint-based approaches*, Ph.D. thesis, Stanford University, Stanford, CA, 2009.

---

<sup>2</sup>In practice, we can rule out the first hundred sampling points to be sure that we have reached the attractor  $\Lambda$

- [9] D.S COHAN, A. HAKAMI, Y. HU, A.G. RUSSELL, *Nonlinear response of ozone to emissions: Source apportionment and sensitivity analysis*, Environmental Science & Technology, 39(17), 6739-6748, 2005.
- [10] R.M. HICKS, P.A. HENNE, *Wing design by numerical optimization*, Journal of Aircraft, 15.7: 407-412, 1978.
- [11] G. GALLAVOTTI, *Entropy, thermostats, and chaotic hypothesis*, Chaos: An Interdisciplinary Journal of Nonlinear Science, 16(4):043114-043114, 2006.
- [12] D. RUELLE, *Measures describing a turbulent flow*, Annals of the New York Academy of Sciences, 357(1):1-9, 1980.
- [13] D. LEA, M. ALLEN, T. HAINE, *Sensitivity analysis of the climate of a chaotic system*, Tellus, 52A:523-532, 2000.
- [14] Q. WANG, R. HU, P. BLONIGAN, *Least Squares Shadowing Sensitivity analysis of chaotic limit cycle oscillations*, Journal of Computational Physics, 2014.
- [15] G. EYINK, T. HAINE, D. LEA, *Ruelle's linear response formula, ensemble adjoint schemes and Lévy flights*, Nonlinearity 17, 1867-1889, 2004.
- [16] R.V. ABRAMOV, A.J. MAJDA, *Blended response algorithms for linear fluctuation-dissipation for complex nonlinear dynamical systems*, Nonlinearity, 20(12):2793, 2007.
- [17] F. COOPER, P. HAYNES, *Climate sensitivity via a nonparametric fluctuation-dissipation theorem*, Journal of the Atmospheric Sciences, 68(5):937-953, 2011.
- [18] Q. WANG, *Convergence of the Least Squares Shadowing method for computing derivative of ergodic averages*, SIAM, J. Numer. Anal. 52 (1) 156-170, 2014.
- [19] S.P. KUZNETSOV, *Hyperbolic Chaos: A Physicist's View*, Springer Berlin Heidelberg, 2012.
- [20] C. BONATTI, L. DIAZ, M. VIANA, *Dynamics Beyond Uniform Hyperbolicity : A Global Geometric and Probabilistic Perspective*, Encyclopaedia of Mathematical Sciences, Springer, 2010.
- [21] C. ROBINSON, *Dynamics systems: Stability, Symbolic Dynamics, and Chaos*, CRC Press, Boca Raton, FL, 1998.
- [22] C. ROBINSON, *Structural stability of C1 flows*, Dynamical Systems—Warwick 1974, 262-275, Springer Berlin Heidelberg, 1975.

**Appendix A.** In our case, we have an explicit expression for  $(v_i^{\{h,\infty\}}, \eta_i^{\{h,\infty\}})$  :

$$v_i^{\{h,\infty\}} = \sum_{j=0}^{\infty} (D\varphi_0(u_{i-j}^0, jh)) \partial_s \varphi_0(u_{i-j}^0, h)^- - \sum_{j=1}^{\infty} (D\varphi_0(u_{i+j}^0, -jh)) \partial_s \varphi_0(u_{i+j}^0, h)^+ \quad (\text{A.1})$$

$$\eta_i^{\{h,\infty\}} = -\frac{1}{h} \frac{\langle \partial_s \varphi_0(u_i^0, h)^0; \partial_h \varphi_0(u_i^0, h) \rangle}{\langle \partial_h \varphi_0(u_i^0, h); \partial_h \varphi_0(u_i^0, h) \rangle} \quad (\text{A.2})$$

As we did previously:

$$\|v_i^{\{h,\infty\}}\| \leq \sum_{j=0}^{\infty} \|(D\varphi_0^{(j)}(u_{i-j}^0, h)) \partial_s \varphi_0(u_{i-j}^0, h)^-\| + \sum_{j=1}^{\infty} \|(D\varphi_0^{(-j)}(u_{i+j}^0, h)) \partial_s \varphi_0(u_{i+j}^0, h)^+\| \quad (\text{A.3})$$

$$\leq \sum_{j=0}^{\infty} C\lambda^{jh} \|\partial_s \varphi_0(u_{i-j}^0, h)^-\| + \sum_{j=1}^{\infty} C\lambda^{jh} \|\partial_s \varphi_0(u_{i+j}^0, h)^+\| \quad (\text{A.4})$$

$$\leq \frac{2C}{1-\lambda^h} \frac{\sup_{u \in \Lambda} (\|\partial_s \varphi_0(u, h)\|)}{\beta} \quad (\text{A.5})$$

so :

$$\|v_i^{\{h,\infty\}}\| \leq \frac{2C}{1-\lambda^h} \frac{\sup_{u \in \Lambda} (\|\partial_s \varphi_0(u, h)\|)}{\beta} \leq \frac{2C}{\beta \ln(\frac{\epsilon}{\lambda})} \frac{\sup_{u \in \Lambda} (\|\partial_s \varphi_0(u, h)\|)}{h} \quad (\text{A.6})$$

$$\leq \frac{2C}{\beta \ln(\frac{\epsilon}{\lambda})} \sup_{u \in \Lambda} \left( \left\| \frac{\partial_s \varphi_0(u, h) - \partial_s \varphi_0(u, 0)}{h} \right\| \right) \quad (\text{A.7})$$

$$\leq \frac{2C}{\beta \ln(\frac{\epsilon}{\lambda})} \sup_{(u,h) \in \Lambda \times H} (\|\partial_h \partial_s \varphi_0(u, h)\|) = \|\mathbf{v}^{\{\infty\}}\| \quad (\text{A.8})$$

where  $H$  is a compact set of  $\mathbb{R}^+$  (for example  $[0, 1]$ ) and  $\partial_s \varphi_0(u, 0) = 0$ . We've also used the taylor expansion of  $1 - \lambda^h$  for  $h \rightarrow 0$  and assumed that  $\partial_h \partial_s$  is well-defined and continuous on the compact set  $\Lambda \times H$ .

Following similar steps, we obtain :

$$\eta_i^{\{h,\infty\}} \leq \frac{\sup_{u \in \Lambda} (\|\partial_s \varphi_0(u, h)\|)}{\beta h m} \quad (\text{A.9})$$

$$\leq \frac{\sup_{(u,h) \in \Lambda \times H} (\|\partial_h \partial_s \varphi_0(u, h)\|)}{\beta m} = \|\boldsymbol{\eta}^{\{\infty\}}\| \quad (\text{A.10})$$

**Appendix B.** Let  $C \in \mathbb{R}^+$  be an arbitrary bound and we will assume that  $\alpha = 1$  for simplicity reasons (without loss of generality). If  $n_e$  is the number of elements that are bigger or equal to  $C$ , we have:

$$n_e \leq \frac{T(\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \|\boldsymbol{\eta}^{\{\infty\}}\|^2)}{hC^2} \quad (\text{B.1})$$

with the equality being verified in the worst case scenario where all the unbounded terms are equal to  $C$ , all the bounded terms are equal to 0 and  $n_e C^2$  is exactly equal to  $\frac{T}{h}(\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \|\boldsymbol{\eta}^{\{\infty\}}\|^2)$ . Then, let us compute the contribution of the terms that are unbounded. For that purpose, we introduce an indicator function  $\delta$  that is equal to 0 when both  $e_i^{\{h,T\}0}$  and  $\epsilon_i^{\{h,T\}}$  are below  $C$  and equal to 1 when at least one of them is bigger (or equal) than  $C$ . We have:

$$\lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \delta(i) \left[ (DJ(u_i, s)) e_i^{\{h,T\}0} + (\epsilon_i^{\{h,T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right) \quad (\text{B.2})$$

$$\leq \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \delta(i) \left[ \|DJ\|_{\infty}^{\Lambda} \|e_i^{\{h,T\}0}\| + (2\|\epsilon_i^{\{h,T\}}\| \|J\|_{\infty}^{\Lambda}) \right] \right) \quad (\text{B.3})$$

$$\leq \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \max(\|DJ\|_{\infty}^{\Lambda}, 2\|J\|_{\infty}^{\Lambda}) \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \delta(i) (\|e_i^{\{h,T\}0}\| + \|\epsilon_i^{\{h,T\}}\|) \right) \quad (\text{B.4})$$

$$\leq \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \max(\|DJ\|_{\infty}^{\Lambda}, 2\|J\|_{\infty}^{\Lambda}) \times n_e C \right) \quad (\text{B.5})$$

$$\leq \lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \max(\|DJ\|_{\infty}^{\Lambda}, 2\|J\|_{\infty}^{\Lambda}) \times \frac{T(\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \|\boldsymbol{\eta}^{\{\infty\}}\|^2)}{hC^2} C \right) \quad (\text{B.6})$$

$$\leq \frac{\max(\|DJ\|_{\infty}^{\Lambda}, 2\|J\|_{\infty}^{\Lambda}) (\|\mathbf{v}^{\{\infty\}}\|_{\mathbb{B}}^2 + \|\boldsymbol{\eta}^{\{\infty\}}\|^2)}{C} \quad (\text{B.7})$$

We have used the fact that  $\sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} \delta(i) (\|e_i^{\{h,T\}0}\| + \|\epsilon_i^{\{h,T\}}\|)$  is maximized when we have the maximum number of unbounded elements and when all of them have the same value. The worst case scenario is again the one where we have  $n_e$  unbounded elements all equal to  $C$ . Since  $C$  is arbitrarily big, the contribution of the unbounded terms is as small as we want. Then, the limits can be permuted in the expression :

$$\lim_{h \rightarrow 0} \lim_{T \rightarrow \infty} \left( \frac{h}{T} \sum_{i=1}^{\lfloor \frac{T}{h} \rfloor} (1 - \delta(i)) \left[ (DJ(u_i, s)) e_i^{\{h,T\}0} + (\epsilon_i^{\{h,T\}} (J(u_i, s) - \langle J \rangle(s))) \right] \right) \quad (\text{B.8})$$

$$= \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( (1 - \delta(i)) \frac{h}{T} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{(J(u_i^s, s) - J(u_i^s, s))}{s} \right) \quad (\text{B.9})$$

$$+ \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( (1 - \delta(i)) \frac{h \epsilon_i^{\{h,T\}}}{T} (J(u_i^s, s) - \frac{1}{T} \sum_j h J(u_i^s, s)) \right) \quad (\text{B.10})$$

$$= \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \lim_{s \rightarrow 0} \left( (1 - \delta(i)) \frac{1}{s} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{h(\tau_i^s + s \epsilon_i^{\{h,T\}}) J(u_i^s, s)}{h \sum_j (\tau_j^s + s \epsilon_j^{\{h,T\}})} - \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} \right) + O(s) \quad (\text{B.11})$$

$$= \lim_{s \rightarrow 0} \lim_{h \rightarrow 0} \lim_{T \rightarrow +\infty} \left( (1 - \delta(i)) \frac{1}{s} \sum_{i=0}^{\lfloor \frac{T}{h} \rfloor} \frac{h(\tau_i^s + s \epsilon_i^{\{h,T\}}) J(u_i^s, s)}{h \sum_j (\tau_j^s + s \epsilon_j^{\{h,T\}})} - \frac{h \tau_i^s J(u_i^s, s)}{h \sum_j \tau_j^s} \right) + O(s) \quad (\text{B.12})$$

$$= \lim_{s \rightarrow 0} \frac{\langle J \rangle(s) - \langle J \rangle(s)}{s} + O(s) = 0 \quad (\text{B.13})$$

The last equality comes from property (4.5) we had on Riemann sums.